

## A BAYESIAN METHOD FOR THE ANALYSIS OF THE DUST EMISSION IN THE FAR-INFRARED AND SUBMILLIMETER

M. VENEZIANI<sup>1,2</sup>, F. PIACENTINI<sup>2</sup>, A. NORIEGA-CRESPO<sup>1</sup>, S. CAREY<sup>1</sup>, R. PALADINI<sup>1</sup>, AND D. PARADIS<sup>3,4</sup>

<sup>1</sup> Infrared Processing and Analysis Center, California Institute of Technology, Pasadena, CA 91125, USA; [marcella.veneziani@ipac.caltech.edu](mailto:marcella.veneziani@ipac.caltech.edu)

<sup>2</sup> Dipartimento di Fisica, Università di Roma “La Sapienza,” I-00185 Rome, Italy

<sup>3</sup> Université de Toulouse, UPS-OMP, IRAP, F-31062 Toulouse, France

<sup>4</sup> CNRS, IRAP, F-31028 Toulouse, France

Received 2012 May 15; accepted 2013 May 14; published 2013 July 5

### ABSTRACT

We present a method, based on Bayesian statistics, to fit the dust emission parameters in the far-infrared and submillimeter wavelengths. The method estimates the dust temperature and spectral emissivity index, plus their relationship, properly taking into account the statistical and systematic uncertainties. We test it on three sets of simulated sources detectable by the *Herschel Space Observatory* in the PACS and SPIRE spectral bands (70–500  $\mu\text{m}$ ), spanning over a wide range of dust temperatures. The simulated observations are a one-component interstellar medium and two two-component sources, both warm (H II regions) and cold (cold clumps (CCs)). We first define a procedure to identify the better model, then we recover the parameters of the model and measure their physical correlations by means of a Markov Chain Monte Carlo algorithm adopting multi-variate Gaussian priors. In this process, we assess the reliability of the model recovery and of parameter estimation. We conclude that the model and parameters are properly recovered only under certain circumstances and that false models may be derived in some cases. We applied the method to a set of 91 starless CCs in an interarm region of the Galactic plane with low star formation activity, observed by *Herschel* in the Hi-GAL survey. Our results are consistent with a temperature-independent spectral index.

**Key words:** H II regions – infrared: ISM – ISM: clouds – methods: data analysis – methods: statistical – surveys

**Online-only material:** color figures

### 1. INTRODUCTION

Previous observations in the submillimeter and far-infrared domains have enabled the investigation of the properties of dust in a variety of environments and at different angular resolutions. The brightness of radiation emitted from a source in local thermal equilibrium has a continuum spectrum which can be expressed as  $I_\lambda = \epsilon_\lambda B_\lambda(T)$ , where  $\epsilon_\lambda$  is the wavelength-dependent emissivity and  $B_\lambda(T)$  is the Planck function corresponding to a temperature  $T$ . In optically thin regions, the emissivity equals the optical depth  $\tau_\lambda$ . In optically thick regions the emissivity tends to 1. In the case of optically thin dust clouds (Désert et al. 1990; Draine & Li 2007), the simplest models assume that the emissivity depends on the wavelength as a power law, and the brightness spectrum is

$$I_\lambda(\epsilon_0, \beta, T_d) = \epsilon_0 \left( \frac{\lambda}{\lambda_0} \right)^{-\beta} B_\lambda(T_d), \quad (1)$$

where  $T_d$  is the dust temperature,  $\epsilon_0$  is the emissivity at wavelength  $\lambda_0$ , and  $\beta$  is the emissivity spectral index, with a fiducial value  $\beta = 2$ .

More refined models describe the spectrum as due to a combination of components different in temperature and nature. In particular, observational data evidence a flattening of the spectral index toward long wavelengths ( $\lambda > 500 \mu\text{m}$ ) which is well represented by a multi-component dust model described in Finkbeiner et al. (1999). The spectrum is then given by

$$I_\lambda(\epsilon_1, \epsilon_2, T_1, T_2) = \epsilon_1 \left( \frac{\lambda}{\lambda_0} \right)^{-\beta_1} B_\lambda(T_1) + \epsilon_2 \left( \frac{\lambda}{\lambda_0} \right)^{-\beta_2} B_\lambda(T_2), \quad (2)$$

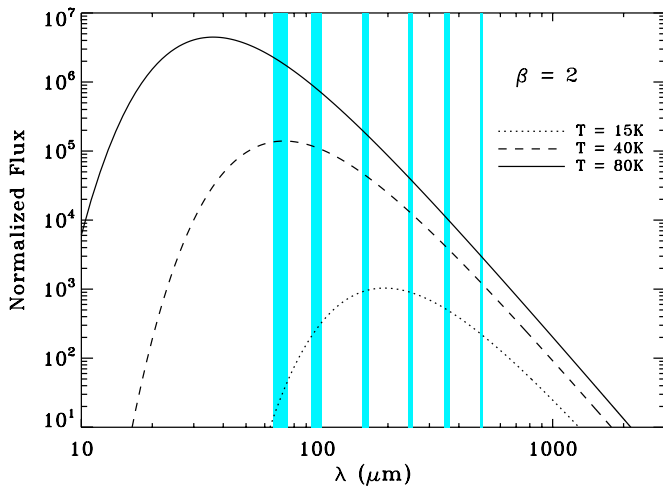
where  $\epsilon_1$  and  $\epsilon_2$  are the emissivities at  $\lambda_0$ ,  $T_1$  and  $T_2$  are the temperatures of the two dust components, and  $\beta_1$  and  $\beta_2$  are the spectral indices. Standard value for spectral indices can be set to  $\beta_1 = 1.67$ ,  $\beta_2 = 2.70$ , i.e., the best-fit values in model 8 of Finkbeiner et al. (1999).

More sophisticated models incorporate the effect of the disordered internal structure of amorphous dust grains (Mény et al. 2007). In this case, the emission is characterized by a single temperature and a spectral index which changes with temperature. Moreover, some level of anticorrelation between spectral index and temperature is expected by laboratory experiments (see, for example, Coupeaud et al. 2011). The emission can then be characterized by Equation (1) with the extra relation

$$\beta = A \left( \frac{T_d}{T_0} \right)^\alpha, \quad (3)$$

where  $T_0$  is a pivot temperature (we use  $T_0 = 20 \text{ K}$ ), and  $A$  and  $\alpha$  are parameters that characterize the inverse relation.

Measurements from balloon-based observatories and satellites indicate the existence of an anticorrelation between  $\beta$  and  $T_d$ . The data sets from the balloon-borne experiments PRONAOS (Dupac et al. 2003) and ARCHEOPS (Désert et al. 2008) evidenced an inverse relationship between  $T_d$  and  $\beta$  in a wide range of environments of the interstellar medium (ISM) and in cold sources. In PRONAOS data, variations in the spectral index were observed in the range 0.8–2.4 for dust temperatures between 11 and 80 K, whereas ARCHEOPS data showed that the inverse relationship is more pronounced with values of  $\beta$  up to 4 for temperatures down to 7 K. Moreover, recently, Veneziani et al. (2010) highlighted a similar trend analyzing  $T_d$  and  $\beta$  for eight high Galactic latitudes clouds, by combining *IRAS*, *DIRBE*, and *Wilkinson*



**Figure 1.** Normalized SEDs for a 15 K, 40 K, and 80 K source and spectral index  $\beta = 2$ . The blue vertical lines identify the *Herschel* bands chosen to perform our analysis. The SEDs are then well constrained at low temperatures (dotted line) and poorly constrained at high temperatures (solid line) where *Herschel* sampling covers only the RJ part of the spectrum.

(A color version of this figure is available in the online journal.)

*Microwave Anisotropy Probe* data to BOOMERanG observations. The  $\beta$  values vary from 1 to 5 in the temperature range 7–20 K, with a behavior similar to the ARCHEOPS results. The analysis of the *Herschel* Hi-Gal key program (Molinari et al. 2010a, 2010b) has demonstrated this anticorrelation in the ISM (Paradis et al. 2010) while the *Herschel*-ATLAS survey found similar results at high latitudes (Bracco et al. 2011). On the contrary, Paladini et al. (2012) find the same inverse trend on H II regions but it might be generated from spurious effects.

In order to test the hypothesis of the existence of a functional dependency between the temperature and the spectral index, one has to properly account for spurious correlations in their estimated values. Due to the spectral shape in Equation (1), the effect of a high value of  $T_d$  can be mimicked, when holding the intensity fixed, by a low value of  $\beta$ , and vice versa. There exists a statistical degeneracy between these physical quantities that must be considered while investigating if an intrinsic physical correlation indeed exists. Shetty et al. (2009a, 2009b) report spurious inverse correlations due to a number of factors including noise in flux measurements coupling with the  $T_d$ – $\beta$  degeneracy and variations in temperature of overlapping sources along the sight line.

Because of the last *Herschel* and *Planck* submillimeter and millimeter data on dust, the topic of the  $T_d$ – $\beta$  spurious anticorrelation has been recently addressed (see, for example, Juvela & Ysard 2012a, 2012b; Kelly et al. 2012) both on simulations and on real data. The originality of our method, based on Bayesian statistics, is to treat statistical and systematic uncertainties in two different ways, both during the spectral energy distribution (SED) fitting and in the  $T_d$ – $\beta$  relationship estimate, taking into account their different statistical properties. This allows us to distinguish between a spurious and a physical anticorrelation between the spectral index and the temperature, and to estimate the  $T_d$ – $\beta$  inverse relationship by properly evaluating the degeneracy caused by the spectral shape and the flux uncertainties, both statistical and systematic, i.e., calibration uncertainty.

We demonstrate the robustness of this method in the *Herschel* PACS and SPIRE spectral coverage,  $70 \mu\text{m} < \lambda < 500 \mu\text{m}$ , and for a temperature range  $10 < T < 50$  K. We consider three cases: the diffuse ISM, warm sources (H II regions/young stellar objects (YSOs)), and cold sources (e.g., pre-stellar clumps).

This paper is organized as follows: Section 2 presents the algorithm, after a brief introduction on the temperature–spectral index degeneracy and Bayesian statistics, Section 3 describes the simulated data set used in the analysis, and Sections 4 and 5 report the results obtained by applying the method to the simulated observations and to cold clumps (CCs) on the Galactic plane detected by the *Herschel* Hi-GAL survey. Conclusions are drawn in Section 6.

## 2. BAYESIAN METHOD FOR IR SED FITTING

In this section, after a brief discussion of the temperature–spectral index degeneracy and of Bayesian statistics, we present our method to fit the SEDs taking into account both the statistical noise and the systematic uncertainties present in every data set. The combination of Bayesian techniques and a good knowledge of the parameter distribution in the parameter space, allows to remove the spurious temperature–spectral index anticorrelation, as well as the biases on the physical parameter estimates, generated both from random noise and systematic effects.

### 2.1. Bayesian Treatment of Statistical Noise

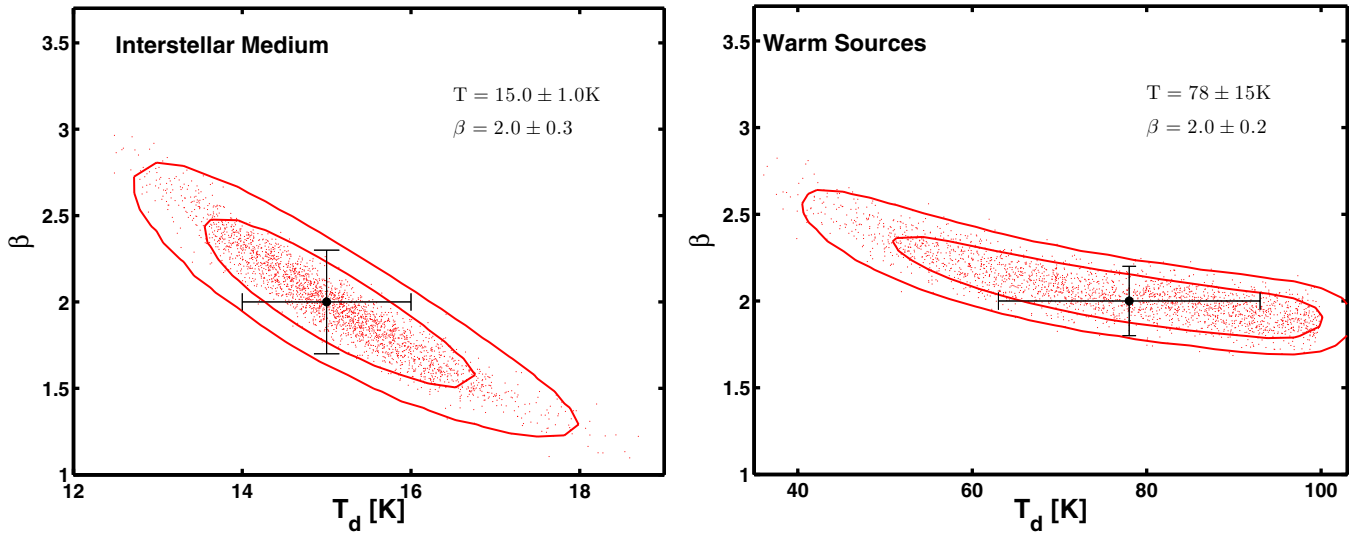
In order to study the effect of noise (i.e., statistical uncertainties) on one-component SED fitting, and in particular the  $T_d$ – $\beta$  degeneracy, we first simulate one ISM-like and one H II-region-like source with a 20% Gaussian random noise, and then fit for the physical parameters using Equation (1). Since we are interested in *Herschel*-like observations, we sample the SEDs in the six bands  $\lambda = [70, 100, 160, 250, 350, 500] \mu\text{m}$ . The simulated source temperatures are  $T_{\text{ISM}} = 15$  K and  $T_{\text{H II}} = 80$  K. The reason for choosing such different temperatures is shown in Figure 1 where the normalized SED of three sources at 15 K, 40 K, and 80 K and their sampling with the *Herschel* bands are reported. In the *Herschel* bands, the SED is well measured at low temperature (dotted line) while for warmer sources (dashed and solid lines) we can only sample the Rayleigh–Jeans (RJ) part of the spectrum without any information on the Wien side. The spectral index is constrained by the slope of the RJ portion of the modified blackbody while the temperature is constrained by the peak location. In our simulation, we fix the spectral index to 2. The  $70 \mu\text{m}$  band is usually not included in ISM studies because it is very sensitive to very small grain (VSG) emission. We include it in this simulation to show the dependence of the parameter recovery on the error bars and on the SED sampling.

To fit the SED of the simulated sources and to study the parameter distribution in the parameter space, as well as their correlation, we use a Markov Chain Monte Carlo (MCMC) algorithm Lewis & Bridle (2002). Hereafter, we explain the main idea behind this approach.

Following Bayesian statistics, the probability of having a set of parameters  $\mathbf{p}$  given a set of data  $\mathbf{d}$  is

$$P(\mathbf{p}|\mathbf{d}) \propto P(\mathbf{p})P(\mathbf{d}|\mathbf{p}). \quad (4)$$

Here  $P(\cdot|\cdot)$  are conditional probability densities with the argument on the right-hand side of the bar assumed to be true,  $P(\mathbf{p})$  is the a priori probability density, or prior, of the parameters set  $\mathbf{p}$ ,  $P(\mathbf{d}|\mathbf{p})$  (e.g., the likelihood function) is the probability



**Figure 2.** Sixty-eight and ninety-five percent contours of the two-dimensional posterior probabilities of sources at  $T = 15$  K,  $\beta = 2$  (left panel) and  $T = 80$  K,  $\beta = 2$  with a 20% statistical error on fluxes. The independent points of the Markov chains are denser within the 68% contour, close to the maximum of the distribution, than close to the boundaries. The slant shape means that  $T$  and  $\beta$  are degenerate and the elongation is larger for the warm sources where *Herschel* bands do not constrain the peak of the SEDs. The  $1\sigma$  error bars are obtained by marginalizing the tridimensional posterior probability, as we have three parameters, over the remaining two dimensions.

(A color version of this figure is available in the online journal.)

density of data set  $\mathbf{d}$  given a set of parameters  $\mathbf{p}$  and is described by the formula

$$P(\mathbf{d}|\mathbf{p}) \propto \exp \left( -\frac{1}{2} \sum_{b=1}^{N_\lambda} \left( \frac{d_b - I_b(\mathbf{p})}{\sigma_b} \right)^2 \right), \quad (5)$$

where  $N_\lambda$  is the number of bands,  $I_b$  is the value of the chosen model with parameters  $\mathbf{p}$  in the  $\lambda_b$  band, either Equation (1) or Equation (2), and  $\sigma_\lambda$  is the statistical error associated with the flux  $d_b$ . If Equation (1) is chosen, then  $\mathbf{p} = [\epsilon_0, T_d, \beta]$ ; otherwise, if Equation (2) is chosen, then  $\mathbf{p} = [\epsilon_1, \epsilon_2, T_1, T_2]$ . The a priori probability density of a given set of parameters is the probability distribution that expresses the parameter uncertainties before the data are taken into account. When nothing is known about the parameter distributions, usually wide flat priors are adopted, as they assign equal probabilities to all possibilities. If, on the contrary, the parameter statistical distribution is known from previous experience, i.e., previous measurements, then a more specific, informative prior might be adopted.

The posterior probability  $P(\mathbf{p}|\mathbf{d})$  is estimated using the MCMC algorithm. Given a set  $\mathbf{p}_i$ , with likelihood  $L_i$  and posterior probability  $P(\mathbf{p}_i|\mathbf{d})$ , the MCMC algorithm performs a random walk through the allowed parameter space, generating a new independent set  $\mathbf{p}_{i+1}$  with likelihood  $L_{i+1}$  and posterior probability  $P(\mathbf{p}_{i+1}|\mathbf{d})$ . This second set is accepted according to a rule which also guarantees a good sampling of the probability density in a reasonable computational time. We make use of the Metropolis–Hastings algorithm, which is a class of Markov chains in which the new set  $\mathbf{p}_{i+1}$  is always accepted if

$$\Lambda(i+1, i) = \frac{P(\mathbf{p}_{i+1}|\mathbf{d})}{P(\mathbf{p}_i|\mathbf{d})} = \frac{L_{i+1}P(\mathbf{p}_{i+1})}{L_iP(\mathbf{p}_i)} > 1. \quad (6)$$

This guarantees the convergence to the maximum of the likelihood function. If instead  $\Lambda(i+1, i) < 1$ , the new set is accepted with a probability proportional to the ratio  $\Lambda(i+1, i)$ . An advantage of this method is to ensure a good sampling of the

parameter distribution in parameter space. For an application of this technique on millimeter and submillimeter observations, see Veneziani et al. (2010).

Figure 2 shows the 68% and 95% posterior probabilities in the  $T_d$ – $\beta$  parameter space obtained for the two cases of ISM ( $T_{\text{ISM}} = 15$  K,  $\beta = 2$ ) and H II region ( $T_{\text{H II}} = 80$  K,  $\beta = 2$ ), respectively. The elongated and slant shape of these posterior probabilities indicate that the two parameters are degenerate in parameter space. In the case of warm sources (right panel), the RJ side of the SED is well sampled by the *Herschel* band and  $\beta$  is therefore well constrained. On the contrary, the location of the peak is uncertain, leading to a large error in the temperature determination.

The dependence of one of the two parameters on the other, i.e., their statistical degeneracy, can be easily demonstrated. Starting from a one-temperature modified blackbody:

$$F(T_b, \beta) = \epsilon_0 \left( \frac{\lambda}{\lambda_0} \right)^{-\beta} \frac{2hc^2}{\lambda^5} \frac{1}{e^{\frac{hc}{\lambda k T_d}} - 1} \Delta\Omega. \quad (7)$$

In the RJ regime, we can approximate with

$$F(T_b, \beta) \simeq \epsilon_0 \left( \frac{\lambda}{\lambda_0} \right)^{-\beta} \frac{2ckT_d}{\lambda^4} \Delta\Omega \quad (8)$$

which is a linear relation in the  $\log\lambda$ – $\log F$  scale. Thus, we expect some level of anticorrelation between  $T_d$  and  $\beta$ , as in any linear relation there is some level of anticorrelation between slope and intercept.

The more the fluxes are noisy and the SED not well sampled, the more the two parameters are correlated and difficult to constrain. Moreover, when combining data of different sources to recover correlation between parameters, there is the tendency to treat systematic errors as if they were statistical, which is not correct. Here, we treat the systematic errors in a proper way, and use Bayesian statistics, in order to provide a correct reconstruction of the parameter distributions and of physical correlations between parameters.

**Table 1**  
Priors for the Model Identification (Step 1)

	<b>p</b>	Range of Variability	<b>P(p)</b>
Model 1	$\epsilon_0$	$>0$	U
	$\beta$	0.5–4	U
	$T_d$ (K)	5–100	U
Model 2	$\epsilon_1$	$>0$	U
	$\epsilon_2$	$>0$	U
	$T_1$ (K)	5–25	U
	$T_2$ (K)	10–100	U
	$\beta_C$	1.7, 2.0, 2.3, 2.7	U

**Notes.** List of a priori probability densities imposed on parameter set during the model identification procedure. We chose wide uniform ranges (U) in order not to constrain the fit results and to avoid the code to converge to nonphysical values. The explored range of  $\beta_C$  values is also reported.

## 2.2. Bayesian Treatment of Systematic Errors

The sources simulated in Section 2.1 have up to 20% of statistical uncertainty on measured fluxes. Nonetheless, the flux uncertainty might come also from the calibration error and other sources of systematic error, which do not have the same properties as the statistical errors (i.e., instrumental noise, background fluctuations). In order to properly take into account the systematic effects, we make use of a Monte Carlo procedure. In the following, we will refer only to the systematic error due to calibration uncertainty but the same procedure can be applied to any source of systematic error.

We would like to emphasize that this treatment is particularly important in our specific case: the recovery of a relationship among physical parameters of different observables, several sources in our case. The systematic error can be the same in all sources, as in the case of a wrong calibration of one band. If this is not considered, the result can be severely biased.

In order to properly take into account in our analysis the calibration error, we use a *conditional inference* technique. We consider an  $n$ -tuple of the calibration values  $\mathbf{k} = (k_1, k_2, \dots, k_{N_k})$ . Each value  $k_b$  has a probability distribution  $P(k_b)$ , which we assume to be uniform in the  $-\Delta k_b < k_b < \Delta k_b$  range. Given the observed data set  $\mathbf{d}$ , we have a joint distribution of the parameters  $\mathbf{p}$  for all possible configurations of  $\mathbf{k}$ ,  $P(\mathbf{p}|\mathbf{d}, \mathbf{k})$ . The conditional result for a single source is obtained by marginalization of this probability over all the possible values of the

calibration uncertainties, weighted by their distribution:

$$P(\mathbf{p}|\mathbf{d}) = \int_{-\Delta k}^{\Delta k} P(\mathbf{d}, \mathbf{k}|\mathbf{p}) P(\mathbf{p}) P(\mathbf{k}) d\mathbf{k}, \quad (9)$$

where

$$P(\mathbf{d}, \mathbf{k}|\mathbf{p}) \propto \exp \left( -\frac{1}{2} \sum_{b=1}^{N_k} \left( \frac{d_b - k_b I_b(\mathbf{p})}{\sigma_b} \right)^2 \right). \quad (10)$$

Equation (9) is an extension of Equation (4) where both statistical and systematic uncertainties are taken into account. In order to apply this technique, we run a Monte Carlo simulation, scattering the values of  $\mathbf{k}$  in 100 realizations. We tried different numbers of calibration realizations, but we could not find any difference in the final results above 100 steps. In each iteration  $\mu$ , a calibration uncertainty vector  $\mathbf{k}_\mu$  is generated. The fluxes  $\mathbf{d}$  of all sources are multiplied by the same set of  $\mathbf{k}_\mu$ . The fit of the SED in each iteration is performed with an MCMC algorithm (see Section 2.1).

The final physical parameters which better describe the SED of each source are then obtained by marginalizing the 100 calibration dependent set of physical parameters over the calibration uncertainties as in Equation (9). The marginalization is performed with the publicly available GetDist software (Lewis & Bridle 2002).

## 2.3. Method Description and Analysis of Uncertainties

The method to recover dust physical parameters focuses on our treatment of statistical and systematic uncertainties and on a good definition of the a priori probability density in the MCMC. The analysis is carried out according to the following steps. The details of each step are reported later in this section.

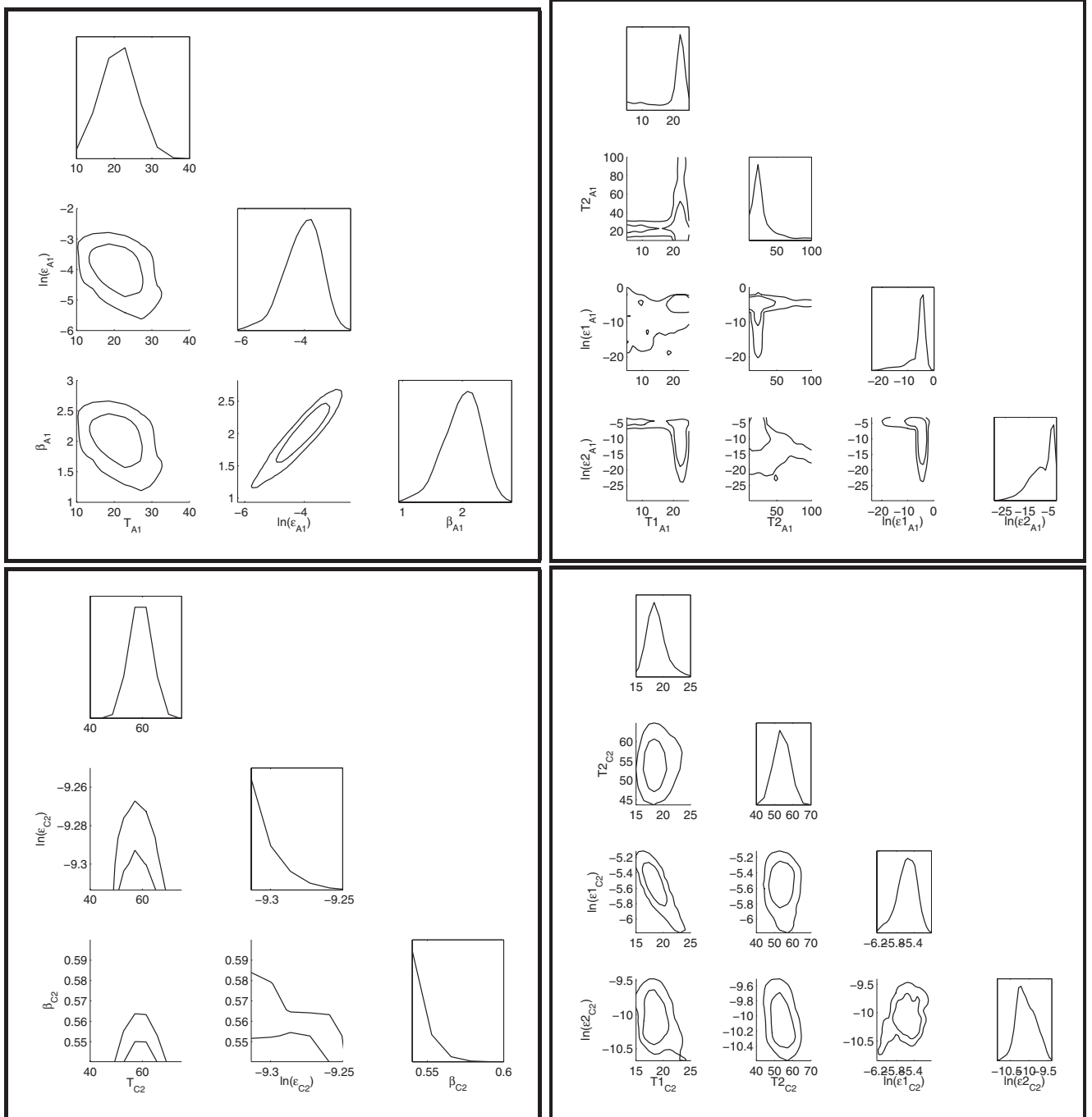
1. We make a first, fast MCMC run on the whole sample of sources to check the best model for each SED and to study the parameter distribution in the parameters space. Therefore, we fit the SED of each source with both Equations (1) and (2) (hereafter  $M_1$  and  $M_2$ , respectively, taking into account both statistical and systematic uncertainties (i.e., calibration errors) as described in Section 2.2. The spectral indices  $\beta_1$  and  $\beta_2$  for  $M_2$  are set to a common value  $\beta_C$ . At this stage, we assume uniform wide priors, reported in Table 1.
2. We determine which model,  $M_1$  or  $M_2$ , better fits the SED of each source, based on the comparison of the  $\chi^2$  probability density functions, as reported in Section 2.4; if both models

**Table 2**  
Simulated Observations

Band ( $\mu\text{m}$ )	Experiment	Beam Size	$\Delta k_{\text{ext}}$	$\Delta k_{\text{point}}$	$\sigma$
24	MIPS (Carey et al. 2009)	6''	15%	4%	10%
70	PACS (Poglitsch et al. 2010)	6''	20%	3%	10%
100	IRAS-IRIS (Miville-Deschênes & Lagache 2005)	4'	13.5%	13.5%	10%
160	PACS (Poglitsch et al. 2010)	11''	20%	4%	10%
250	SPIRE (Griffin et al. 2010)	18''	15%	7%	10%
350	SPIRE (Griffin et al. 2010)	25''	15%	7%	10%
500	SPIRE (Griffin et al. 2010)	37''	15%	7%	10%
850	Planck-HFI (Planck Collaboration 2011)	5'	3%	3%	10%

**Notes.** Experimental characteristics used to generate the simulated sources. The reported calibration uncertainties ( $\Delta k_{\text{ext}}$  and  $\Delta k_{\text{point}}$ ) refer to diffuse emission and to point sources, respectively. The statistical error  $\sigma$  is assumed to be 10%, which is the maximum value obtained in the analysis of *Herschel* data in the Hi-GAL survey. It includes both instrumental noise and background fluctuations.





**Figure 3.** One- and two-dimensional physical parameters posterior probabilities for an ISM-like (source A1, see Table 2) and an H II (source C2) region-like source estimated with an MCMC run assuming wide uniform priors. The ISM is well fitted by  $M_1$  (top left) and not by  $M_2$  (top right). As a consequence, the posterior probabilities are well constrained within the range provided when considering a one-temperature model (top left) and not with a two-temperature model (top right). The opposite is true for the second one:  $M_1$  is not well constrained (bottom left) and converges to very low value of the spectral index  $\beta_{HII}$ , while all the unknowns in  $M_2$  converge to finite results (bottom right).

give good probability, we assign the source the simpler  $M_1$ ; if both models give bad probability we exclude the source from the analysis. An example of good and bad fits after step 1 is reported in Figure 3.

3. We fit again the physical parameters of the sources classified as  $M_1$  with updated prior information. Therefore, we assume multi-variate Gaussian priors, estimated from the first run. This procedure is described in Section 2.5.
4. From the new physical values, the temperature–spectral index relationship is measured as reported in Section 2.6.

## 2.4. Model Identification

In real observations, especially of the Galactic plane, there is often an overlap of sources along the line of sight or a multi-component temperature source. It is crucial to be able to distinguish among different models and identify the correct one, before attempting any physical interpretation.

As a first step in our analysis, we then fit the simulated data set with both models in Equations (1) and (2), i.e., with a single temperature and two temperatures. In  $M_1$ , the fitted parameters are the emissivity ( $\epsilon_0$ ), the temperature  $T_d$ , and the

**Table 3**  
Simulated Sources

Input Sources	Central Temperature (K)	Relative Column Densities	$\beta$	Resolution	Spectral Range ( $\mu\text{m}$ )
Interstellar medium (A1)	$T_A = 20$		2	4'	100–500
Interstellar medium (A2)	$T_A = 20$		$2.2(T_A/T_0)^{-1.3}$	4'	100–500
Interstellar medium (A3)	$T_A = 20$		$2(T_A/T_0)^{+0.5}$	4'	100–500
Cold clump (B1)	$T_C = 10, T_E = 15, T_{\text{col}} = 12.5$	$N_C = N_E$	2	5'	100–500
Cold clump (B2)	$T_C = 10, T_E = 15, T_{\text{col}} = 12.5$	$N_C = N_E$	2	5'	100–850
Cold clump (B3)	$T_C = 10, T_E = 15, T_{\text{col}} = 10.9$	$N_C = 10 N_E$	2	5'	100–500
Cold clump (B4)	$T_C = 10, T_E = 15, T_{\text{col}} = 10.9$	$N_C = 10 N_E$	2	5'	100–850
H II+dust (C1)	$T_{\text{H II}} = 17, T_{\text{dd}} = 50, T_{\text{col}} = 17.3$	$N_{\text{dd}} = 100 N_{\text{H II}}$	2	4'	70–500
H II+dust (C2)	$T_{\text{H II}} = 17, T_{\text{dd}} = 50, T_{\text{col}} = 17.3$	$N_{\text{dd}} = 100 N_{\text{H II}}$	2	4'	24–500

**Notes.** Three types of simulated sources used to test our method. In each model, we generate 1000 sources with a Gaussian distribution in temperature. The distribution is contained on the value reported in the second column, with a dispersion of 10%. In cases A2 and A3, there is a  $T_d$ – $\beta$  correlation in input.  $T_0$  is set to 20 K.

spectral index  $\beta$ . We do not consider any model with more than two components because with the data sets available nowadays it is unlikely to have enough data points to well fit up to six parameters. For the same reason, in the two-component case, we cannot perform a fit over all six parameters:  $T_1$ ,  $T_2$ ,  $\epsilon_1$ ,  $\epsilon_2$ ,  $\beta_1$ , and  $\beta_2$ . We rather set the spectral indices to a fixed common value  $\beta_c = \beta_1 = \beta_2$ , and estimate the two emissivities and the two temperatures. The values explored for  $\beta_c$  in model 2 are 1.7, 2.0, 2.3, and 2.7, consistently with Finkbeiner et al. (1999). We then execute the fit for  $M_2$  four times, changing every time the spectral index which is assumed to be the same for both the components, i.e., not temperature dependent. To make sure not to constrain the results, we assume uniform wide a priori probability densities on every parameter. They are the same for all sources and are reported in Table 1.

We can then easily find the model which better fits the SED by using the  $\chi^2$  probability density function. Since  $M_1$  and  $M_2$  have different degrees of freedom, we compare the two analyses using the cumulative distribution function  $P$  of a  $\chi^2$  distribution with the proper number of degrees of freedom. The ideal fit should have  $P \sim 0.5$ .  $P \ll 0.5$  indicates an overestimate of the error bars while  $P \gg 0.5$  indicates that the fitting model is not a good model for the data set. In one-component sources, we expect, in an ideal situation,  $P(M_1)$  to be near to 0.5, and  $P(M_2)$  to be larger. The opposite is true for two-component sources. In an ideal situation, we expect to have at least one of the  $P(M_2; \beta_c)$  near 0.5, and  $P(M_1)$  to be larger.

An example of one- and two-dimensional posterior probabilities of the physical parameters of an ISM-like source and of an H II region combined with dust is shown in Figure 3. In the top line of this figure, we show the posterior probabilities of a 20 K source (ISM) fitted with a one-temperature model (top left panel) and with a two-temperature model (top right panel). The convergence is reached with the one-temperature model because all the three parameters (temperature, spectral index, and emissivity) are well constrained within the boundary provided. On the contrary, the fit with two components (top right) is not well constrained. The four parameters (two temperatures and two emissivities) do not converge within the acceptable ranges. In the bottom line of Figure 3, the fit of a two-component source with one- and two-component models is shown. The source is an H II region of 50 K combined with some surrounding dust at 17 K. The spectral indices of both components are set to 2. In this case, the fit with a two-component model converges (bottom right) with the four parameters well defined within the chosen ranges, while the fit with a one-component model (bot-

tom left) is poorly constrained within the physically acceptable boundaries.

### 2.5. Parameter Estimates

Since the parameters in Equations (1) and (2) are not uniformly distributed and are also intrinsically correlated in the parameter space, their “a priori” probability density  $P(\mathbf{p})$  (Equation (9)) must take these properties into account. This is required to avoid biased estimates of the source characteristics. The intrinsic parameter correlation in each source is due to the spectral shape describing the source emission (Equations (1) and (2)) and it does not have anything to do with the physical dependency of one parameter on the other, which can be measured from an ensemble of objects. The level of the intrinsic correlation, in our case, depends on the random noise present in the single SEDs, on the SED sampling within the considered bands and on possible source overlap along the line of sight. Since those effects might vary source by source, we analyze the parameter statistical distribution of each source independently of the others. The information about the parameter statistical distribution and intrinsic correlation is contained in the covariance matrix, estimated for each source by means of the first MCMC run (step 1). The covariance matrix is then used to build the new a priori probability distribution, which is not wide and flat as in the first run, but contains specific and definite information about the variables. The new set of priors  $P(\mathbf{p})$  is a multi-variate Gaussian, with the center  $\bar{\mathbf{p}} = [\log \epsilon_0, \beta_0, T_0]$  in the approximate average values of the whole sample of sources. The covariances of the multi-variate Gaussian priors are estimated, for each source, by broadening 10 times the parameters standard deviations keeping their correlations. This makes us sure not to constrain the final results with a too tight parameters range. In other words, while the center of the Gaussian prior is the same for the whole set of sources, the parameter covariances are estimated source by source and, as a consequence, the fit is performed source by source.

Taking all those aspects into account, the final parameters estimate on each source is then performed again by applying Equations (9) and (10) but, this time, instead of assuming uniform distributed priors, we describe  $P(\mathbf{p})$  as a multi-variate Gaussian:

$$P(\mathbf{p}) \propto \exp \left( -\frac{1}{2} ((\mathbf{p} - \bar{\mathbf{p}})^T \Sigma^{-1} (\mathbf{p} - \bar{\mathbf{p}})) \right), \quad (11)$$

where  $\Sigma_{ij}$  is the covariance matrix of each source, with the indices  $i$  and  $j$  running over the entire set of parameters.

**Table 4**  
Priors during the Final Estimate (Step 3)

Sources	$\mathbf{p}$	$\bar{\mathbf{p}}$	$\sigma_{\bar{\mathbf{p}}}$	$C$	$P(\mathbf{p})$
ISM (A)	$\log \epsilon_0$	-4.2	9.7	$C(\log \epsilon_0, \beta) = 0.985$	MVG
	$\beta$	1.8	2.5	$C(\log \epsilon_0, T_d) = -0.985$	MVG
	$T_d$ (K)	23	85.7	$C(\beta, T_d) = -0.956$	MVG
CC (B)	$\log \epsilon_0$	-6.5	6.1	$C(\log \epsilon_0, \beta) = 0.982$	MVG
	$\beta$	1.8	1.6	$C(\log \epsilon_0, T_d) = -0.986$	MVG
	$T_d$ (K)	13	54.8	$C(\beta, T_d) = -0.951$	MVG
H II+ISM (C)	$\log \epsilon_0$	-5.5	34.9	$C(\log \epsilon_0, \beta) = 0.969$	MVG
	$\beta$	1.8	9.0	$C(\log \epsilon_0, T_d) = -0.976$	MVG
	$T_d$ (K)	23	309.3	$C(\beta, T_d) = -0.924$	MVG

**Notes.** List of a priori probability densities imposed on the final parameter estimates. The covariance matrix ( $\Sigma$ ) elements are obtained by multiplying the correlation matrix elements ( $C$ ) with the standard deviation of the two correspondent parameters. The values reported are for just one source for each kind of simulated observations. The overall shape of the probability density is a multi-variate Gaussian (MVG) where we only take the positive values.

$\Sigma_{ij}$  elements are obtained from the correlation matrix  $C_{ij}$  by  $\Sigma_{ij} = C_{ij} \sigma_i \sigma_j$ , where  $\sigma_i \sigma_j$  are the standard deviations of the parameters  $p_i$  and  $p_j$ , obtained from the first MCMC run, multiplied by a factor of 10 in order to broaden the prior, retaining the correlation properties.

An example of  $C$  elements is reported in Table 4.  $\Sigma$  contains therefore the information about the distribution and correlation of parameters. Both  $\bar{\mathbf{p}}$  and  $\Sigma$  are estimated through the MCMC in step 1. The final likelihood is then

$$P(\mathbf{p}|\mathbf{d}) \propto \int_{-\Delta\mathbf{k}}^{\Delta\mathbf{k}} \exp\left(-\frac{1}{2} \sum_{b=1}^{N_s} \left(\frac{d - kI(\mathbf{p})}{\sigma_d}\right)_b^2\right) \cdot \exp\left(-\frac{1}{2} \sum_{i=1}^{N_p} \sum_{j=1}^{N_p} \left(\frac{(p_i - \bar{p}_i)C^{-1}(p_j - \bar{p}_j)}{\sigma_i \sigma_j}\right)\right) d\mathbf{k}, \quad (12)$$

where  $N_p$  is the number of parameters and  $\mathbf{k}$  is the vector of the band calibration uncertainties on extended emission. As already pointed out, with our method sources are fitted one by one since parameters of different objects might be correlated in a different way due to random noise or different systematic effects, for example, SED sampling or overlap of other sources along the line of sight.

If the sources in the sample have a similar noise level and systematic characteristics, i.e., line-of-sight temperature variations and calibration uncertainties, one can also use a hierarchical Bayesian procedure (Kelly et al. 2012). The hierarchical model fits the SEDs of the whole sample of sources at the same time, assuming a global model for the overall distribution of source parameters.

### 2.6. Estimate of the Temperature–Spectral Index Relationship

As discussed in Section 1, an inverse relation between temperature and spectral index has been found in several previous studies. We model this relation by means of Equation (3).

The comparison between  $M_1$  and  $M_2$  and the estimate of the functional form of the parameter distributions, are the first steps to estimate the temperature–spectral index relationship. We select as possible source candidates of a temperature-dependent spectral index, the ones with  $M_1$  preferred over  $M_2$  and with

$P(M_1) < 0.95(2\sigma)$ . This last requirement is necessary because the worse the fit, the more spurious the  $T_d$ – $\beta$  anticorrelation.

If all the previous conditions are satisfied, we estimate the  $T_d$ – $\beta$  trend on the selected subsample of sources, taking into account the calibration uncertainties again through a Monte Carlo procedure. In each iteration  $\mu$ , i.e., for each  $\mathbf{k}_\mu$ , we fit all sources with  $M_1$  and estimate Equation (3) with an MCMC run assuming uniform wide priors on  $A_\mu$  and  $\alpha_\mu$ :  $0 < A < 3$ ,  $-2 < \alpha < 2$ . After 100 iterations, we have 100 pairs of  $A_\mu$  and  $\alpha_\mu$ . We then marginalize over the calibration errors as in Equation (9):

$$P(A, \alpha|\mathbf{x}) = \int_{-\Delta\mathbf{k}}^{\Delta\mathbf{k}} P(\mathbf{x}; \mathbf{k}|A, \alpha) P(A, \alpha) P(\mathbf{k}) d\mathbf{k}, \quad (13)$$

where  $\mathbf{x} = (T_d, \beta)$ ,  $P(\mathbf{k})$  is the same as in Equation (9),  $P(A, \alpha)$  is assumed to be uniformly distributed and uncorrelated, and the likelihood of the  $\mu$ th iteration is

$$P(\mathbf{x}; \mathbf{k}|A, \alpha) \propto \exp\left(-\frac{1}{2} \sum_{s=1}^{N_s} \left(\frac{\beta(\mathbf{k}_\mu) - A_\mu \left[\frac{T_d(\mathbf{k}_\mu)}{T_0}\right]^{\alpha_\mu}}{\sigma(\beta(\mathbf{k}_\mu))}\right)_s^2\right), \quad (14)$$

where  $N_s$  is the number of sources,  $\beta(\mathbf{k}_\mu)$  and  $T_d(\mathbf{k}_\mu)$  are the spectral index and temperature of source  $s$ , estimated with the calibration uncertainties  $\mathbf{k}_\mu$ , and  $\sigma(\beta(\mathbf{k}_\mu))$  is the statistical error on  $\beta(\mathbf{k})$ . One-dimensional probabilities of  $A$  and  $\alpha$  are obtained by marginalization over the other parameter.

## 3. SIMULATED OBSERVATIONS

Since we are interested in studying the SED fitting using the *Herschel* bands, as a baseline of our analysis we use the PACS and SPIRE wavelengths, angular resolutions, and noise properties. We complement the *Herschel* spectral coverage also simulating fluxes in the MIPS (24  $\mu\text{m}$ ), *IRAS* (100  $\mu\text{m}$ ), and *Planck*-HFI (850  $\mu\text{m}$ ) bands. These bands are particularly helpful to constrain the SEDs of warm and cold single and multi-temperatures objects whose emission is not fully sampled by the PACS and SPIRE instruments. Depending on the model, we simulate a one- or two-temperature modified blackbody with a  $1\sigma$  statistical uncertainty of  $\sigma_\lambda = 10\%$  of the total flux in each band. This value arises from the combination of the experimental *Herschel* noise confusion limit (few percent of the total flux) combined with some non-negligible background fluctuation and it is the value also chosen by Shetty et al. (2009a) in their analysis. The calibration uncertainties are estimated on diffuse emission. The fluxes are then scattered within the error bars twice: first, in a Gaussian uncorrelated fashion, to take into account random noise; second, in a correlated fashion, i.e., the same bands have the same scatter in all sources, to take into account the calibration uncertainties. We assume a common angular resolution, i.e.,  $4'$ , for *IRAS* 100  $\mu\text{m}$ , or  $5'$  when *Planck* 850  $\mu\text{m}$  is also included. The considered bands and the experimental characteristics used to perform the simulations are reported in Table 2.

We simulate observations of sources with different physical conditions. The simplest case consists of one-component sources, described by Equation (1), with a temperature around 20 K and a spectral index which may or may not vary with temperature (cases A1 and A2/A3, respectively). A more complex case is given by the overlap of more than one source along the line of sight. Indeed, one of the key questions of the SED fitting is the error made when approximating a multi-temperature

observation with an isothermal model and if this approximation can lead to a spurious correlation among the fitted parameters (Shetty et al. 2009b). This could be the case when analyzing a source with more than one population of dust grains or more than one source along the line of sight. We have considered two different scenarios, one in which the two temperatures are very close (few degrees apart) and the other in which they are very different (tens of degrees). A typical source of the first case is a cold dense pre-stellar clump, i.e., a CC with a temperature ( $T_C$ ) around 10 K surrounded by an envelope of dust with an average temperature ( $T_E$ ) of around 15 K (case B). A typical source of the second case might be an H II region combined with foreground/background diffuse dust (case C). Both the considered scenarios are described by Equation (2).

These last two cases are also chosen to investigate the effect of the band coverage. A good sampling of the SED peak and of the RJ side is required, in order to properly fit both temperature and spectral index and to identify the underlying physical model. For example, in H II regions, the *Herschel* wavelength coverage is not enough to separate the two components, while including the MIPS 24  $\mu\text{m}$  the two components are properly detected. Similarly, pre-stellar clumps properties are better identified including *Planck*-HFI 850  $\mu\text{m}$  band.

The simulated sources are described below and more details are provided in Table 3.

- [A] *ISM environment, as in model 1.* Single-component sources with a Gaussian temperature distribution centered in 20 K and ranging between 15 K and 25 K, observed with a 4' resolution in the spectral range 100–500  $\mu\text{m}$ . We place ourselves exactly in the same situation as Paradis et al. (2010) in order to test our ability to recover the real relationship between temperature and spectral index. For this purpose, we first set the spectral index to 2 (sources A1) and then we use as input a temperature-dependent spectral index, both anticorrelated, as in Paradis et al. (2010), (sources A2) and positively correlated (sources A3).
- [B] *CCs, based on model 2.* Two-component sources consisting of an envelope at a temperature  $T_E$  and a core at a temperature  $T_C$ . The spectral indices of both components are set to 2. We investigate two options for the relative emissivities as analyzed by Shetty et al. (2009b). We also study the dependence of the fit on the spectral coverage by measuring the emission in the range 100–500  $\mu\text{m}$  and then including also the 850  $\mu\text{m}$  flux which constrains the RJ side of the SEDs.
- [C] *H II regions combined with foreground/background diffuse dust, of temperatures  $T_{\text{H II}}$  and  $T_{\text{dd}}$ , respectively.* This case is also based on model 2. The spectral indices are set to 2 and the relative emissivities are scaled according to the analysis performed on Hi-GAL data by Paladini et al. (2012). As in the case of the CCs, we investigate the dependence of the physical parameters recovery on the spectral coverage. We then sample the SEDs in the 70–500  $\mu\text{m}$  range and then include also the 24  $\mu\text{m}$  flux to better constrain the temperature of the warmer component. In this analysis, we assume the 24  $\mu\text{m}$  emission to be dominated by big grains (BGs) as recently found, for example, by SOFIA (Salgado et al. 2012). The BG thermalizes with the interstellar radiation field and we can model their emission as a modified blackbody. If, on the contrary, the 24  $\mu\text{m}$  flux is dominated by VSGs, we need to remove this point from the analysis and consider only the 70–500  $\mu\text{m}$  range,

because they do not have the same physical properties as BG.

In order to compare the temperatures obtained fitting with  $M_1$  the SEDs of two-temperature sources with the input values, we make use of a density-weighted temperature introduced by Doty & Palotti (2002),  $T_{\text{col}}$ . According to their definition, the column temperature of a source described by Equation (2) is given by

$$T_{\text{col}} = \frac{\epsilon_1 T_1 + \epsilon_2 T_2}{\epsilon_1 + \epsilon_2}. \quad (15)$$

#### 4. RESULTS

In this section, we present the results of the SED fitting, and the  $T_d$ – $\beta$  relationship estimate after properly taking into account the systematic errors of the data sets, for each type of source considered.

The multi-variate Gaussian priors of one sample source for each set are reported in Table 4 and shown in Figure 4.

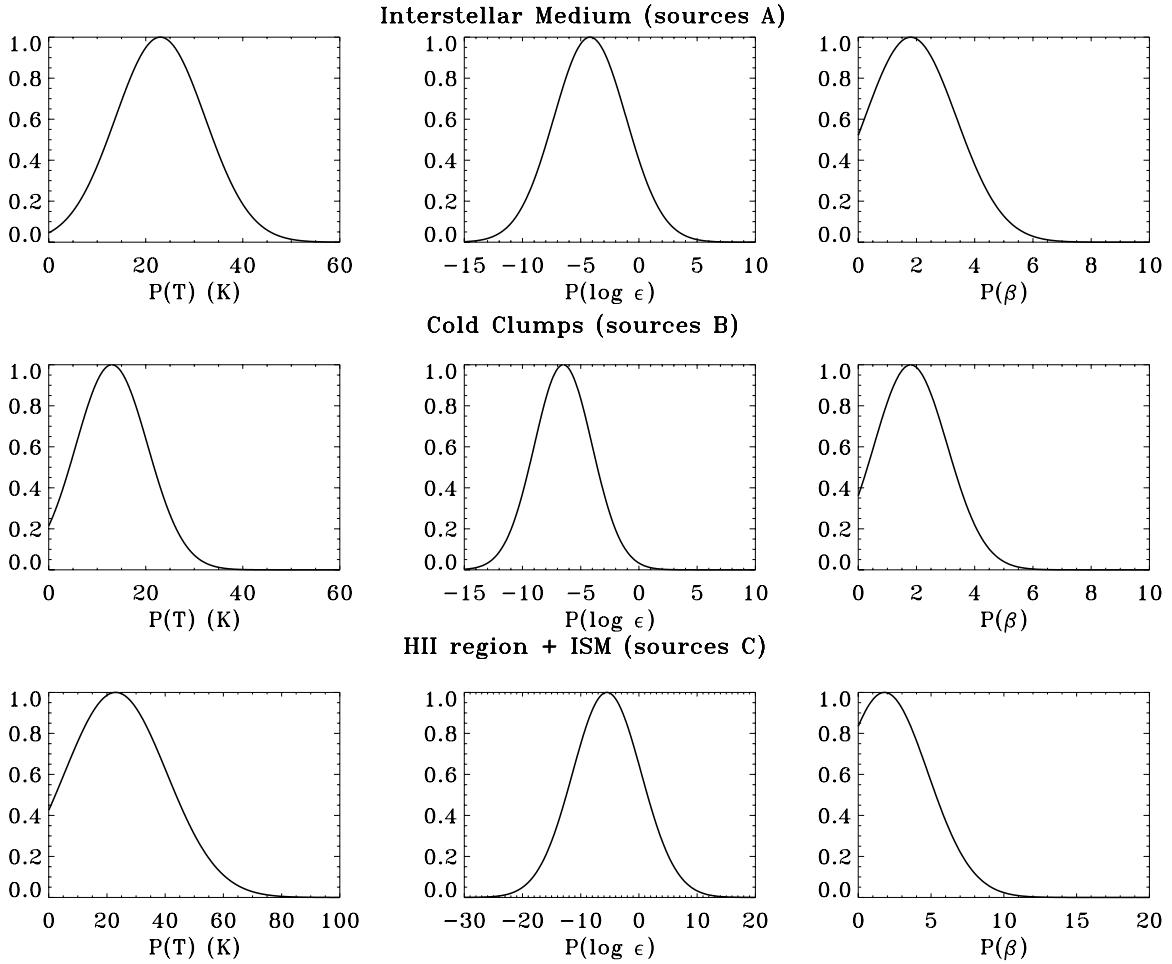
The results are summarized in Tables 5 and 6. In Table 5, first column, we provide the input models; in the second, third, and fourth columns, the percent fraction of sources assigned  $M_1$  and  $M_2$  as well as the excluded sources, respectively, highlighting the true positive detections; columns five to eight report for which  $\beta_c$  value the best  $M_2$  fit is obtained. In Table 6, we give the input and measured  $T_d$ – $\beta$  relationship obtained with our procedure (“Bayesian output”) and with a least-squares method (“least-squares output”), for comparison, for the sources assigned  $M_1$ . The least-squares method is based on a  $\chi^2$  minimization to recover the source parameters and their  $1\sigma$  errors. Since this method does not perceive any treatment of the systematic uncertainties, besides adding them in quadrature to the statistical errors, it has been tested on simulations where only 10% Gaussian error bars are present, without including the additional calibration errors. The least-squares minimization is used in both stages: the SED fitting and the  $T$ – $\beta$  relationship estimate. In this last step, the statistical uncertainties adopted on data points are the ones obtained from the SED fitting step. In order to claim a physical anticorrelation with the Bayesian procedure, we require a  $3\sigma$  detection of the  $\alpha$  parameter.

As a general comment, we would like to highlight that the two-component model is not well constrained if the spectral coverage is poor. In order to have a clear detection of more than one component we need to have information both on the RJ and the Wien part of the SEDs. If the second component is very faint or the spectral coverage too small to clearly detect the second peak, the marginalized final parameters obtained with  $M_2$  are badly constrained.

##### 4.1. Interstellar Medium

The considered cases (A1, A2, and A3) are shown in Figure 5 and reported in Table 5. Their SEDs are much better approximated by  $M_1$  than by  $M_2$ , the two-component model being very badly constrained when the Monte Carlo analysis of the calibration uncertainties is included. As the plots of the relative differences show (panels (a)–(f)), the temperatures are systematically shifted toward higher values of  $\sim 5\%$ – $10\%$  and this corresponds to a systematic spectral index shift toward lower values of  $\sim 10\%$ – $15\%$  of the input on average. In order to understand the origin of this bias, we run five new sets of simulations of the A1 model, all with 10% statistical errors but with [0%, 5%, 10%, 15%, 20%] calibration uncertainties, assuming all the bands to have the same systematic uncertainty. Since we do not





**Figure 4.** One-dimensional projection of the multi-variate Gaussian prior  $P(\mathbf{p})$  described in Equation (11) for each set of sources. The parameters are reported in Table 4.

**Table 5**  
Compared Analysis

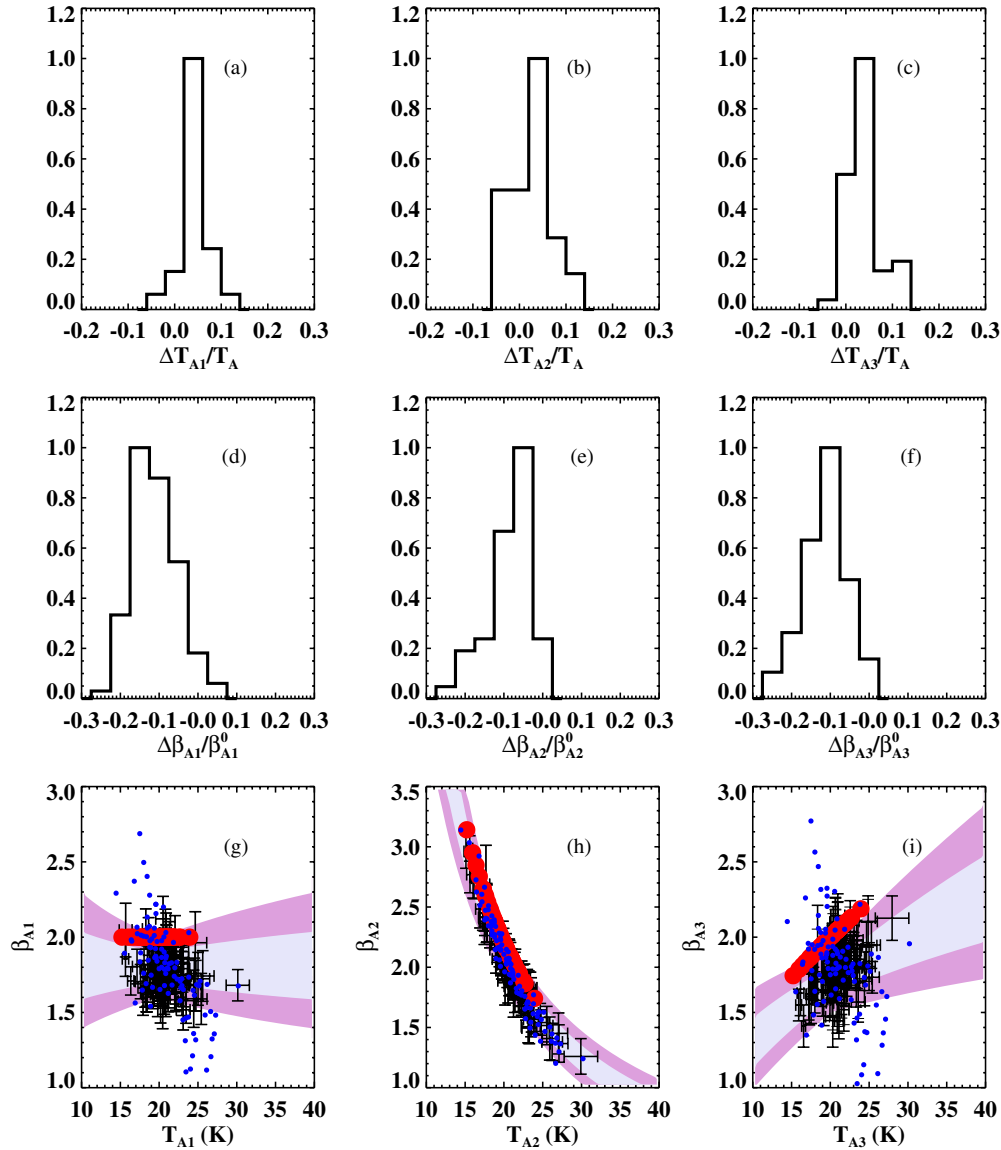
Input	$M_1$ (%)	$M_2$ (%)	Excluded (%)	$\beta_c = 1.7$ (%)	$\beta_c = 2.0$ (%)	$\beta_c = 2.3$ (%)	$\beta_c = 2.7$ (%)
Interstellar medium (A1)	<b>100</b>	0	0	0	0	0	0
Interstellar medium (A2)	<b>98</b>	2	0	2	0	0	0
Cold clump (B1)	100	<b>0</b>	0	0	0	0	0
Cold clump (B2)	99	<b>1</b>	0	0	0	1	0
Cold clump (B3)	89	<b>11</b>	0	0	1	9	1
Cold clump (B4)	86	<b>14</b>	0	0	0	14	0
H II+dust (C1)	32	<b>68</b>	0	1	48	19	0
H II+dust (C2)	4	<b>96</b>	0	9	33	54	0

**Notes.** Comparison of the different kind of sources approximated with a one-component ( $M_1$ ) and a two-component model ( $M_2$ ). We run model 2 four times setting the spectral index  $\beta_c$  to the values 1.7, 2, 2.3, and 2.7 to explore the whole range indicated by Finkbeiner et al. (1999). The bold face values indicate the percentage corresponding to the right input models.

have a more precise knowledge of the distribution of the calibration errors, we can only assume a uniform flat prior in the considered interval. We run the whole pipeline, and estimate the bias on the final temperatures ( $\Delta T/T_0$ ) and spectral indices ( $\Delta\beta/\beta_0$ ). Results on the bias study are shown in Figure 6. When a uniform wide prior is assumed (left panel), which is the case we apply on simulations and on real observations, the systematic shift increases with the calibration uncertainty, in a non-uniform way for the two parameters: the spectral index seems to be affected more than the temperature, likely because the calibration

variations affect the RJ slope of the SEDs more than the peak position. Nonetheless, if we assume to have a better knowledge of the calibration uncertainty distributions, which is not the case for the observations considered in the present paper, and assume Gaussian priors with standard deviations [0%, 3%, 5%, 7%, 10%], the bias level is very low for both the parameters (the systematic shift is centered in few percents) and is consistent with zero (right panel).

A good measurement of this bias would require a better knowledge of the uncertainty distribution. However, our method



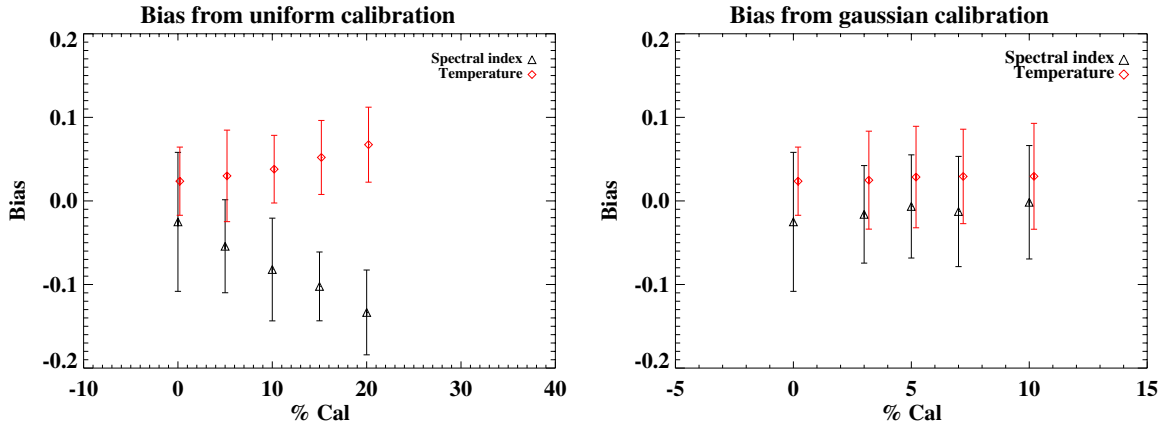
**Figure 5.** Comparison between output and input temperatures and spectral indices for sources A (panels (a)–(f)). All the histograms are normalized to 1. Panels from (g) to (i):  $T$ – $\beta$  relationships. Red circles: input values. Black dots: values recovered with the Bayesian methods, within  $1\sigma$  bars. Blue dots: values recovered with a least-squares method. The light and dark purple contours identify the recovered correlation within  $1\sigma$  and  $2\sigma$  error, respectively.

(A color version of this figure is available in the online journal.)

**Table 6**  
Temperature–Spectral Index Relationship from Simulations

Input Model	Input		Bayesian Output		Input Recovery	Least Square Output		Input Recovery
	$A$	$\alpha$	$A(\mathbf{k}_{\text{ext}}, \sigma_{10\%})$	$\alpha(\mathbf{k}_{\text{ext}}, \sigma_{10\%})$		$A(\sigma_{10\%})$	$\alpha(\sigma_{10\%})$	
Interstellar medium (A1)	2.0	0.0	$1.8 \pm 0.1$	$0.0 \pm 0.1$	True	$1.84 \pm 0.03$	$-0.8 \pm 0.1$	False
Interstellar medium (A2)	2.2	$-1.3$	$2.1 \pm 0.1$	$-1.3 \pm 0.1$	True	$2.01 \pm 0.03$	$-1.4 \pm 0.1$	True
Interstellar medium (A3)	2.0	0.5	$1.7 \pm 0.1$	$0.4 \pm 0.1$	True	$1.83 \pm 0.03$	$-0.5 \pm 0.1$	False
Cold clump (B1)	2.0	0.0	$1.7 \pm 0.1$	$-0.3 \pm 0.1$	True	$1.06 \pm 0.01$	$-1.67 \pm 0.02$	False
Cold clump (B2)	2.0	0.0	$1.7 \pm 0.1$	$-0.3 \pm 0.1$	True	$1.42 \pm 0.04$	$-0.7 \pm 0.1$	False
Cold clump (B3)	2.0	0.0	$2.0 \pm 0.2$	$-0.3 \pm 0.2$	True	$0.5 \pm 0.1$	$-2.3 \pm 0.5$	False
Cold clump (B4)	2.0	0.0	$2.2 \pm 0.2$	$0.0 \pm 0.2$	True	$0.8 \pm 0.1$	$-1.2 \pm 0.3$	False
H II+dust (C1)	2.0	0.0	$1.9 \pm 0.2$	$-0.4 \pm 0.7$	True	$3.0 \pm 0.5$	$-2.4 \pm 0.3$	False
H II+dust (C2)	2.0	0.0	...	...	...			

**Notes.** Output  $T_d$ – $\beta$  relationship (Columns 4–6) obtained with our method, based on Bayesian statistic with multi-variate Gaussian priors, and comparison with the input (Columns 2 and 3) and with the results obtained by fitting the same set of sources with a least-squares method (Columns 7 and 8). Our simulations include a 10% statistical error and a 15%–20% systematic uncertainty depending on the band. Systematic uncertainties are treated only in the procedure based on the Bayesian method. The least-squares method does not allow a correct systematic error treatment, so results based on this procedure include only a 10% statistical uncertainty. For a discussion of these results, see the text.

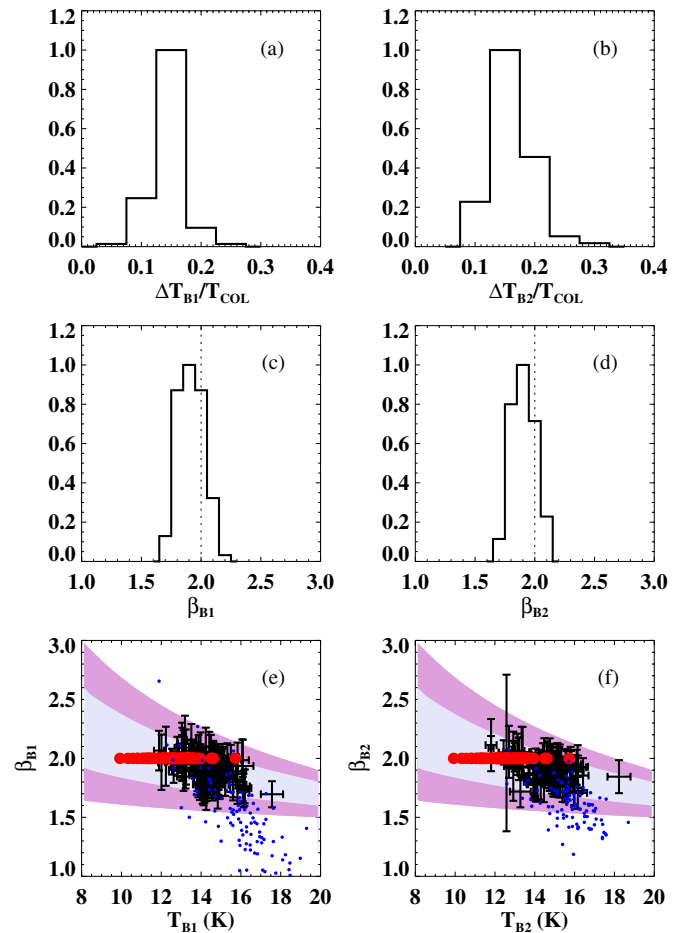


**Figure 6.** Average bias on the final temperatures and spectral indices as a function of the calibration uncertainties.  
(A color version of this figure is available in the online journal.)

takes into account this effect in the estimate of the parameter uncertainties and, therefore, the input values and the recovered  $T-\beta$  are consistent within  $1\sigma$ . As it will be clear in the following, all the parameter estimates in this paper are likely affected by this effect. The systematic shift is slightly more important when more components along the line of sight are assumed (cases B and C) but, again, this effect is included in the final error bars.

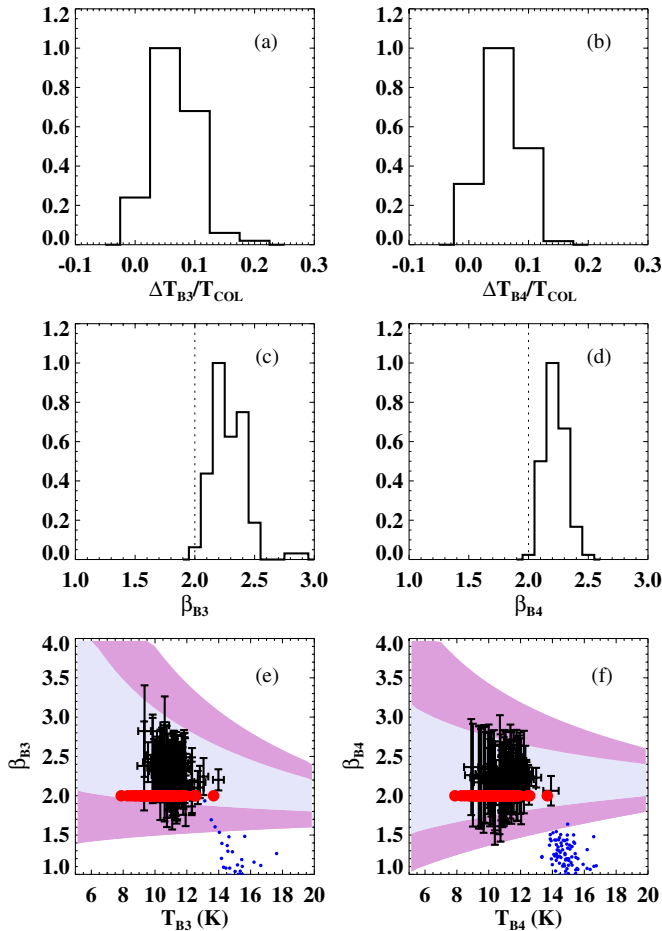
#### 4.2. Cold Clumps

The results of the analysis on CC-type sources are reported in Table 5 (cases B1, B2, B3, and B4) and are shown in Figures 7 and 8. This is the most challenging case considered, as the SEDs are the combination of two modified blackbodies, having temperatures just few degrees apart. It is therefore difficult to identify the presence of two components through the comparison between  $M_1$  and  $M_2$ , especially when the cold component has the same column density as the warm one and is fainter (cases B1 and B2). For these reasons, the majority of sources is assigned to  $M_1$  in those cases, even if they are constituted by two emitting sources. In cases B3 and B4, since the cold component is brighter by construction, it is easier to detect the presence of two sources, even if the majority of sources are still assigned to  $M_1$  and many sources have bad fit. Where the presence of the second component is more evident (cases B3 and B4), the one-component approximation is clearly worst and the parameters are recovered with larger uncertainties. In panels from (a) to (d) of Figure 8, we show the relative differences between the input and output column temperatures and spectral indices for these cases. Cases B1 and B2 are affected from the same systematic bias observed in the ISM simulations, due to large calibration uncertainties. Their parameters are therefore shifted of  $\sim 15\%$  of the original value, on average. This creates also a spurious  $T-\beta$  anticorrelation which is nonetheless negligible within  $3\sigma$ . The same is not true for cases where the cold component is denser and the column density therefore higher and comparable with the column density of the envelope. Here again the parameters, mostly the spectral indices, are affected by some bias due to the wrong modeling spectral shape. Nonetheless, the overall  $T-\beta$  relationship is correct within the uncertainties and is particularly accurate for case B4, where the presence of the  $850\ \mu\text{m}$  point better constraints the RJ slope and, therefore, the spectral index. The  $T-\beta$  relationships are presented in both figures in panels (e) and (f).



**Figure 7.** Results for sources B1 and B2. The comparison between the input and output temperatures (a) and (b) and spectral indices (c) and (d) are shown. The dashed lines mark the input value. Panels (e) and (f): report the input (red circles) and output  $T-\beta$  relationship estimated with our Bayesian method (black dots) and with a least-squares method (blue dots). The light and dark purple contours identify the recovered correlation within  $1\sigma$  and  $2\sigma$  error, respectively.  
(A color version of this figure is available in the online journal.)

An application of the method on a set of cold cores on an interarm region of the Galactic plane observed in the *Herschel* survey is reported in Section 5.



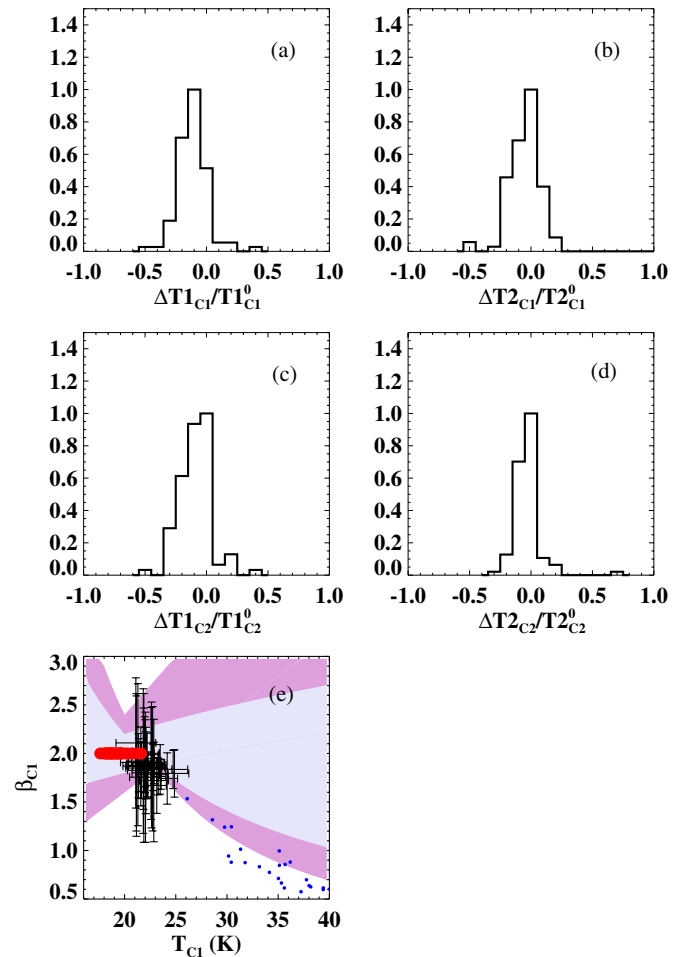
**Figure 8.** Results for sources B3 and B4 (see caption of Figure 7).  
(A color version of this figure is available in the online journal.)

#### 4.3. H II Regions and Dust

Results for H II region-type sources are shown in Figure 9 and reported in Table 5. In both cases, the presence of two components is well detected with the model identification, especially when the  $24\ \mu\text{m}$  flux is present to constrain the peak of the warm component (case C2). As already outlined in Section 3, in this analysis we consider the  $24\ \mu\text{m}$  flux to be dominated by BG. For this kind of sources,  $M_2$  fits better than  $M_1$  in 100% of cases. When excluding the MIPS point (case C1) and we sample the SEDs from  $70\ \mu\text{m}$  down to  $500\ \mu\text{m}$ ,  $\sim 70\%$  of cases are properly classified. Panels (a)–(d) show the correct recovery of the two-component temperatures (ISM and H II) by fitting the SED with  $M_2$ , with a maximum relative error of 50%. They are recovered with a maximum relative error of 50% in the worst cases (panels from (a) to (d)). Panel (e) shows the  $T$ – $\beta$  relationship for sources C1, by fitting the SEDs from  $70\ \mu\text{m}$  to  $500\ \mu\text{m}$  with a modified single blackbody. The  $70\ \mu\text{m}$  flux might still be affected by the warm component emission and this creates an average shift toward warmer temperatures in the parameter recovery. Despite a slight anticorrelation being detected on the data points, it is negligible when the calibration uncertainties are taken into account.

### 5. STARLESS CORES IN THE HERSCHEL/HI-GAL SURVEY

After having tested the validity of the method on simulated data sets, we apply it to a set of starless cores detected in the



**Figure 9.** Comparison between output and input parameters for sources C1 and C2. All the histograms are normalized to 1. Panels from (a) to (d): relative difference on the cold and warm temperatures recovered with  $M_2$  with respect to the input values. Panel (e):  $T$ – $\beta$  relationship from the sources C1 identified as single components. The red circles mark the input column density while the light and dark purple contours identify the recovered correlation within  $1\sigma$  and  $2\sigma$  error, respectively. The same parameters estimated through a least-squares fit are also shown for comparison (blue dots).

(A color version of this figure is available in the online journal.)

*Herschel* Hi-GAL survey (Molinari et al. 2010a, 2010b) in an interarm region of the Galactic plane. Hi-GAL is one of the *Herschel* open time key projects to map the entire Galactic plane in  $2^\circ \times 2^\circ$  tiles with PACS and SPIRE in parallel mode, in the bands 70, 160, 250, 350, and  $500\ \mu\text{m}$ . Once the survey will be complete, the sky coverage will be  $0^\circ < \ell < 360^\circ$  and  $-1^\circ < b < 1^\circ$ . The data set, due to the spectral range, sensitivity, and sky coverage, is particularly sensitive to the very early stages of high-mass star formation. In this paper, we focus on one of the science demonstration phase fields of the survey: a  $2^\circ \times 2^\circ$  tile centered on  $(\ell, b) = (59^\circ, 0^\circ)$ . This field was observed in 2009 November and, since then, it has been largely studied and used to test the analysis algorithms. The tile covers an interarm region, since the line of sight is tangent to the Sagittarius arm, at an heliocentric distance between 2 and 7 kpc, where most of the sources are located. The overall star formation activity is mostly dominated by the Vulpecula OB association (see, for example, Billot et al. 2010), a molecular complex with triggered star-forming activity taking place. Nonetheless, the star formation rate in that region, estimated from the  $70\ \mu\text{m}$  colors of the observed YSOs, is low ( $2.6 \times 10^{-6} M_\odot\ \text{yr}^{-1}$ ;



Veneziani et al. 2013). Since the background emission is also weak, the source detection is very accurate down to very low signal-to-noise thresholds and this makes it a favorable set to study molecular CCs in a pre-stellar evolutionary stage.

### 5.1. Observations and Data Analysis

In this context, we use the data from the *Spitzer* legacy survey MIPS GAL at 24  $\mu\text{m}$  (Carey et al. 2009) and the *Herschel* Hi-GAL data at 70, 160, 250, 350, and 500  $\mu\text{m}$ . Kinematic Sun distances are also estimated from CO observations (Russeil et al. 2011).

*Herschel* maps have been produced by means of the ROMAGAL algorithm, based on a generalized least-squares technique, after a careful preprocessing in which glitches and systematic effects present in the data have been removed. For further information about the pipeline followed from raw data up to sky maps delivery, we refer the reader to Traficante et al. (2011). The source detection and extraction has been performed by means of the CuTEX algorithm (Molinari et al. 2011) which double-differentiates the sky image and study the curvature variations above a given threshold. The identified source profiles are then fitted with a two-dimensional Gaussian plus an underlying inclined planar plateau. The *Herschel* catalog is then produced in the following way: when a source is found in the longest wavelength map (500  $\mu\text{m}$ ), we look for an association in the following band (350  $\mu\text{m}$ ) in a radius as large as the FWHM of the largest beam between the two. If a source is found, then we proceed to the association with the following band (250  $\mu\text{m}$ ) and so on up to the 70  $\mu\text{m}$  map. If more than a source is found in the radius, we associate the nearest. Flux corrections are applied both in the PACS and SPIRE bands in order to take into account the nonperfectly Gaussian shape of the beam. By merging the *Herschel* catalog with the MIPS catalog by means of the same procedure described above, we have an overall catalog of the entire region from 24 to 500  $\mu\text{m}$ .

The procedure for selecting and identifying the starless molecular clumps is summarized as follows.

1. The sources are not detected either in the 70  $\mu\text{m}$  nor in the 24  $\mu\text{m}$  band in order to exclude the presence of a forming star inside the clump.
2. The sources are detected in all the four remaining bands, from 160  $\mu\text{m}$  to 500  $\mu\text{m}$ . This makes us sure not to include spurious detections and to have a better characterization of the SEDs.
3. The sources are gravitationally bound. As in Veneziani et al. (2013), we identify as bounded objects the sources with mass  $M \geq 0.5M_{\text{BE}}$ , where  $M_{\text{BE}}$  is the Bonnor–Ebert (BE) mass. A BE sphere is an isothermal sphere at hydrostatic equilibrium. Therefore, in absence of internal turbulence and assuming thermal pressure, the BE mass is a good approximation of the virial mass.

The absence of turbulence is a necessary approximation due to the fact that no spectroscopic data are available on this sample of sources. We cannot then check the internal motions occurring in the clumps and see if there are supporting mechanisms other than thermal pressure. After all these criteria have been applied we have a total of 103 sources. Since only four fluxes are available, we assign an upper limit of 0.2 Jy to the 70  $\mu\text{m}$  band, in order to perform the model identification. With Hi-GAL sensitivity and coverage, 0.2 Jy is the minimum flux detected in the PACS blue band even at a 15 kpc distance (Veneziani et al. 2013). Twelve sources are excluded through the model identification, leaving

us with a sample of 91 sources. The majority of those objects have radius  $\delta > 0.1$  pc, confirming that they are essentially clumps.

### 5.2. SED Fitting and Dust Parameter Estimate

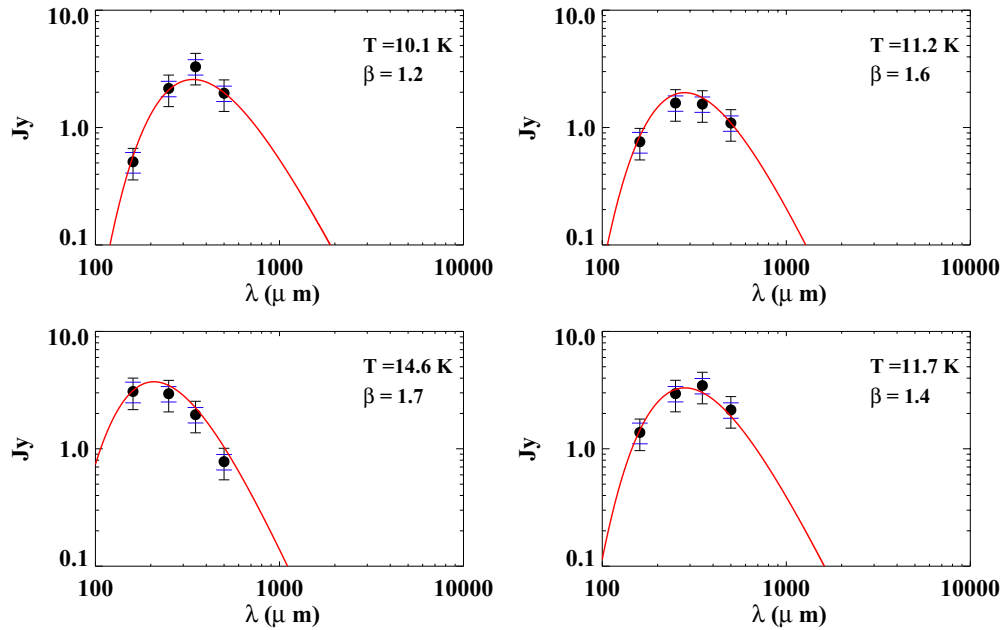
We associate a conservative error bar of 30% of the fluxes with the SEDs from 160  $\mu\text{m}$  to 500  $\mu\text{m}$ . This error takes into account the Gaussian random noise, the source multiplicity, the background fluctuations, and the overall uncertainty on fluxes recovery of the extraction and photometry algorithms. The calibration errors included are PACS and SPIRE uncertainties on extended emission, the same we considered in our simulation, i.e., 20% for the 160  $\mu\text{m}$  PACS band and 15% for the 250, 350, and 500  $\mu\text{m}$  SPIRE bands.

The Bayesian fit with multi-variate Gaussian priors and following marginalization on the calibration uncertainties are then run to estimate the SEDs parameters. After a first MCMC run with uniform wide priors, we fix the center of the multi-variate Gaussian priors in the point  $[T_0, \log(\epsilon_0), \beta_0] = [12, -9, 2]$ . As already described, even if the full sample is collected in the same observation, the covariances are estimated source by source since they depend on random fluctuations, SED sampling, and overlap of other sources along the line of sight, assuming that all the other sources of error, i.e., systematic effects, are the same for all the objects in the same observation. In each covariance matrix, the standard deviations are enlarged by 10 times in order to make sure not to constrain the final parameters with a too tight distribution but keeping the same correlations. The values are nevertheless well described by the ones obtained from the simulation run of cold cores and reported in Table 4.

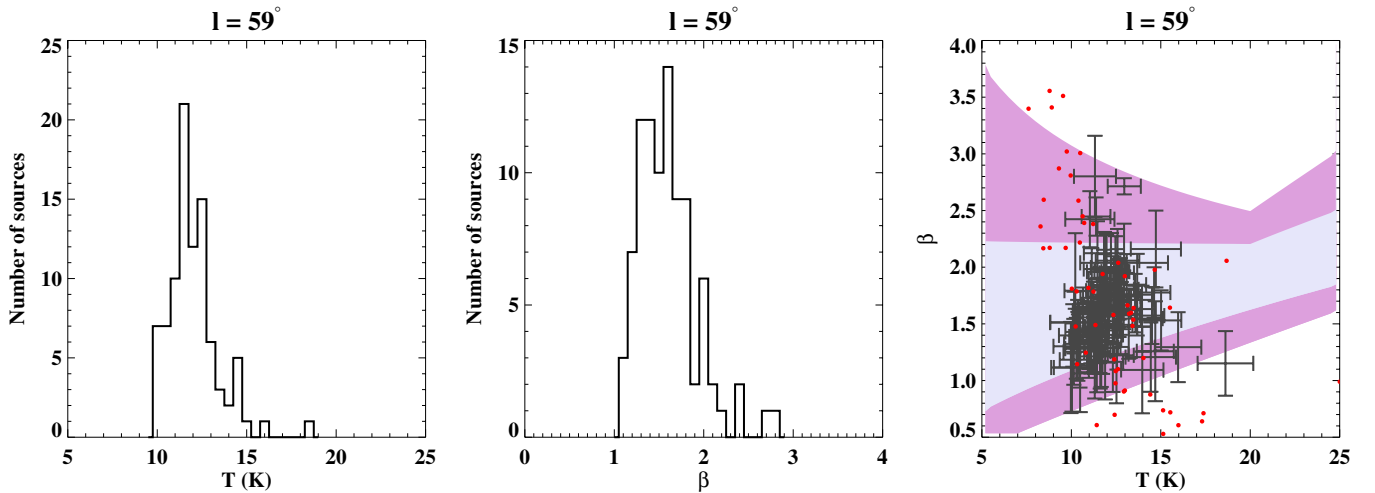
An example of the SEDs with calibration and statistical uncertainties and their best fits is shown in Figure 10.

The final temperature and spectral index distributions, as well as the  $T$ – $\beta$  relationship recovered with our Bayesian method (BM), are shown in Figure 11 as black dots with the associated error bars. Their median values are  $\langle T_{\ell 59} \rangle = 11.8$  K and  $\langle \beta_{\ell 59} \rangle = 1.6$ . The relationship between the temperature and the spectral index has parameter values  $A_{\ell 59}^{\text{BM}} = 1.91 \pm 0.29$  and  $\alpha_{\ell 59}^{\text{BM}} = 0.28 \pm 0.29$ . In order to compare the results of our BM with a technique based on least-squares minimization, we fit the SEDs by associating with the fluxes only a statistical 30% error bar, without taking into account the calibration uncertainties. The temperature and spectral indices values are shown in Figure 11 as red dots. Results from the least-squares fitting show a spurious clear anticorrelation:  $A_{\ell 59}^{\text{LS}} = 0.8 \pm 0.3$  and  $\alpha_{\ell 59}^{\text{LS}} = -1.3 \pm 0.5$  which is, nonetheless, consistent with no anticorrelation within a  $3\sigma$  range.

As already described in Section 1, an increase of the emissivity spectral index in cold environments had been observed both in lab experiments (Mény et al. 2007) and in real observations (Dupac et al. 2003; Désert et al. 2008; Veneziani et al. 2010; Paradis et al. 2010; Bracco et al. 2011). It can be explained in two ways: the spectral index is temperature dependent, as suggested from Mény et al. (2007) from lab experiments on amorphous silicate-based dust grains, or the spectral index increases as a consequence of grain aggregation in cold and dense environments (see, for example, Ormel et al. 2011). In the observed *Herschel* cold cores, we do not detect a significant increase of the spectral index with respect to the ISM, since  $\beta$  values are spread between 1 and 2.5 (Figure 11, middle panel) and, when the  $T$ – $\beta$  trend is studied, points are scattered on the  $T$ – $\beta$  plane excluding the possibility of a correlation (Figure 11, right panel).



**Figure 10.** Example of SEDs of four cold clumps at different temperatures and spectral indices values. The black symbols mark the fluxes with the associated 30% error bars. The blue dashes indicate the calibration uncertainty.  
(A color version of this figure is available in the online journal.)



**Figure 11.** Physical parameters of the starless clumps in the *Herschel*/Hi-GAL  $(\ell, b) = (59^\circ, 0^\circ)$  field. First and second panels show the temperature and spectral index distributions, respectively. The third panel shows the  $T$ - $\beta$  distribution. The light and dark purple contours identify the recovered correlation within  $1\sigma$  and  $2\sigma$  error, respectively. The red dots mark the temperatures and spectral indices obtained by fitting the SEDs with a least-squares method, for comparison.

(A color version of this figure is available in the online journal.)

The absence of a correlation and, more in general, the lack of an increase of the spectral index with respect to the diffuse ISM can be explained considering that, as pointed out several times during the paper, this analysis is very sensitive to systematic effects which might range from instrumental issues other than calibration, (i.e., photometric extraction, catalog band merging), to multiple overlap of sources along the line of sight. A possible physical explanation of our findings might be that we are observing cores of different ages and this reflects the spread of the spectral indices. As Ossenkopf & Henning (1994) show, when silicate dust grains are covered by a thick ice mantle, their emission becomes independent of the optical properties of the material underlying the mantle, as to mask the  $T$ - $\beta$  anticorrelation. Therefore, depending on the mantle thickness, i.e., the timescale of the core, the optical properties of the grains and their spectral indices change. Moreover, as Ormel et al. (2009)

show, depending on the cold core timescale, grains might or might not have the time to aggregate and this also can explain the spread of the spectral index values.

## 6. CONCLUSIONS

We have presented a method, based on Bayesian statistics, to fit for intrinsically correlated parameters and look for the underlying correlation law. This method is different from the ones previously used for similar analysis, as we make use of Bayesian statistics, which fully sample the parameter space taking into account the covariance among different parameters, together with a proper treatment of systematic errors by means of a Monte Carlo procedure. We apply the method to estimate the physical properties and the relationship between the temperature and the spectral index of a set of simulated astrophysical sources detectable in the *Herschel* PACS and SPIRE bands (between

70  $\mu\text{m}$  and 500  $\mu\text{m}$ ), also including *Planck*-HFI (850  $\mu\text{m}$ ), *IRAS* (100  $\mu\text{m}$ ), and MIPS (24  $\mu\text{m}$ ) bands, in order to better constrain the colder and warmer components. The sources are chosen over a wide range of temperatures and compositions in order to test the efficiency of the method with different samplings of the SEDs. For this purpose, we consider the cases of a one-component ISM (with  $T_d$  and  $\beta$  either correlated, cases A2 and A3, or uncorrelated, case A1), and of two-component temperature sources both warm and cold. The case of two-component sources is particularly interesting because, as shown in Shetty et al. (2009a, 2009b), the combination of more than one source along the line of sight generates a spurious anticorrelation by itself.

The most relevant results are as follows.

1. The input physical values and their correlations are well recovered in the *Herschel* bands when we observe an ISM-like sample both without and with correlation (cases A1, A2, and A3, respectively).
2. When two sources just few degrees apart are combined, as CCs, it is difficult to identify the underlying model especially when the core and the envelope have the same column density (cases B1 and B2) and the core emission is therefore much fainter. If the core is thicker it is easier to detect (cases B3 and B4) and, in this case, the presence of the 850  $\mu\text{m}$  flux (B4) helps to better constrain the RJ part of the SEDs and, therefore, the physical parameters and their correlations. Even if a spurious anticorrelation is detected in the CCs, when taking into account the calibration uncertainties, the detected anticorrelation is negligible within  $3\sigma$ .
3. When the source is a combination of two components several degrees apart, as for H II regions, the underlying model is correctly identified, especially when the 24  $\mu\text{m}$  flux is present. Also in this case, a slight anticorrelation is measured in the dust component but the detection is negligible when taking into account calibration uncertainties.

The method has been applied to a sample of starless clumps in a  $2^\circ \times 2^\circ$  field centered on  $(\ell, b) = (59^\circ, 0^\circ)$  observed in the *Herschel* Hi-GAL survey. Being an interarm region of the Galactic plane, it is not very dense with sources and

the star formation activity is therefore very low. The average temperature and spectral index of the sample are  $\sim 11.8$  K and 1.6, respectively, as expected for cold sources. We do not detect a  $T$ - $\beta$  correlation within the uncertainties.

M.V. acknowledges ASI support via contract I/038/080/0. The authors thank the anonymous referee for helpful suggestions that greatly improved the quality of the paper.

## REFERENCES

- Billot, N., Noriega-Crespo, A., Carey, S., et al. 2010, *ApJ*, **712**, 797  
 Bracco, A., Cooray, A., Veneziani, M., et al. 2011, *MNRAS*, **412**, 1151  
 Carey, S. J., Noriega-Crespo, A., Mizuno, D. R., et al. 2009, *PASP*, **121**, 7  
 Coupeaud, A., Demyk, K., Meny, C., et al. 2011, *A&A*, **535**, A124  
 Désert, F.-X., Boulanger, F., & Puget, J. L. 1990, *A&A*, **237**, 215  
 Désert, F.-X., Macías-Pérez, J. F., Mayet, F., et al. 2008, *A&A*, **481**, 411  
 Doty, S. D., & Palotti, M. L. 2002, *MNRAS*, **335**, 993  
 Draine, B. T., & Li, A. 2007, *ApJ*, **657**, 810  
 Dupac, X., Bernard, J.-Ph., Boudet, N., et al. 2003, *A&A*, **404**, L11  
 Finkbeiner, D. P., Davis, M., & Schlegel, D. J. 1999, *ApJ*, **524**, 867  
 Griffin, M. J., Abergel, A., Abreu, A., et al. 2010, *A&A*, **518**, L3  
 Juvela, M., & Ysard, N. 2012a, *A&A*, **539**, 71  
 Juvela, M., & Ysard, N. 2012b, *A&A*, **541**, 33  
 Kelly, B. C., Shetty, R., Stutz, A. M., et al. 2012, *ApJ*, **752**, 55  
 Lewis, A., & Bridle, S. 2002, *PhRvD*, **66**, 103511  
 Mény, C., Gromov, V., Boudet, N., et al. 2007, *A&A*, **468**, 171  
 Miville-Deschênes, M. A., & Lagache, G. 2005, *ApJS*, **157**, 302  
 Molinari, S., Schisano, E., Faustini, F., et al. 2011, *A&A*, **530**, A133  
 Molinari, S., Swinyard, B., Bally, J., et al. 2010a, *PASP*, **122**, 314  
 Molinari, S., Swinyard, B., Bally, J., et al. 2010b, *A&A*, **518L**, 100  
 Ormel, C. W., Min, M., Tielens, A. G. G. M., Dominik, C., & Paszun, D. 2011, *A&A*, **532**, A43  
 Ormel, C. W., Paszun, D., Dominik, C., & Tielens, A. G. G. M. 2009, *A&A*, **502**, 845  
 Ossenkopf, V., & Henning, T. 1994, *A&A*, **291**, 943  
 Paladini, R., Umana, G., Veneziani, M., et al. 2012, *ApJ*, **760**, 149  
 Paradis, D., Veneziani, M., Noriega-Crespo, A., et al. 2010, *A&A*, **520**, L8  
 Planck Collaboration 2011, *A&A*, **536**, A1  
 Poglitsch, A., Waelkens, A., Geis, N., et al. 2010, *A&A*, **518**, L2  
 Russeil, D., Pestalozzi, M., Mottram, J. C., et al. 2011, *A&A*, **526**, A151  
 Salgado, F., Berne, O., Adams, J. D., et al. 2012, *ApJL*, **749**, L21  
 Shetty, R., Kauffmann, J., Schnee, S., & Goodman, A. A. 2009a, *ApJ*, **696**, 676  
 Shetty, R., Kauffmann, J., Schnee, S., Goodman, A. A., & Ercolano, B. 2009b, *ApJ*, **696**, 2234  
 Traficante, A., Calzoletti, L., Veneziani, M., et al. 2011, *MNRAS*, **416**, 2932  
 Veneziani, M., Ade, P. A. R., Bock, J. J., et al. 2010, *ApJ*, **713**, 959  
 Veneziani, M., Elia, D., Noriega-Crespo, A., et al. 2013, *A&A*, **549**, A130